

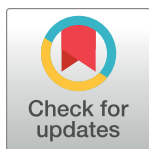
## RESEARCH ARTICLE

# *Burkholderia pseudomallei* distribution in Australasia is linked to paleogeographic and anthropogenic history

Anthony L. Baker<sup>1,2\*</sup>, Talima Pearson<sup>3</sup>, Jason W. Sahl<sup>3</sup>, Crystal Hepp<sup>3,4</sup>, Erin P. Price<sup>5,6</sup>, Derek S. Sarovich<sup>5,6</sup>, Mark Mayo<sup>5</sup>, Apichai Tuanyok<sup>7</sup>, Bart J. Currie<sup>5</sup>, Paul Keim<sup>3</sup>, Jeffrey Warner<sup>2</sup>

**1** Tasmanian Institute of Agriculture (TIA), University of Tasmania, Sandy Bay, Tasmania, Australia, **2** College of Public Health, Medical and Veterinary Sciences, James Cook University, Townsville, Queensland, Australia, **3** Pathogen and Microbiome Institute, Northern Arizona University, Flagstaff, Arizona, United States of America, **4** Informatics and Computing, Northern Arizona University, Flagstaff, Arizona, United States of America, **5** Global and Tropical Health Division, Menzies School of Health Research, Charles Darwin University, Darwin, Northern Territory, Australia, **6** Faculty of Science, Health, Education and Engineering, University of the Sunshine Coast, Sippy Downs, Queensland, Australia, **7** College of Veterinary Medicine, University of Florida, Gainesville, Florida, United States of America

\* [a.l.baker@utas.edu.au](mailto:a.l.baker@utas.edu.au)



## OPEN ACCESS

**Citation:** Baker AL, Pearson T, Sahl JW, Hepp C, Price EP, Sarovich DS, et al. (2018) *Burkholderia pseudomallei* distribution in Australasia is linked to paleogeographic and anthropogenic history. PLoS ONE 13(11): e0206845. <https://doi.org/10.1371/journal.pone.0206845>

**Editor:** Lisa A. Morici, Tulane University School of Medicine, UNITED STATES

**Received:** September 4, 2018

**Accepted:** October 20, 2018

**Published:** November 5, 2018

**Copyright:** © 2018 Baker et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Sequence data generated during this study have been submitted to the Sequence Read Archive database with the following accession numbers: SRR2896252, SRR2896253, SRR2921951, SRR2896254, SRR2896255, SRR2896256, SRR2896271, SRR2896258, SRR2896259, SRR2896260, SRR2896261, SRR2896262, SRR2896263, SRR2896264.

**Funding:** This work was supported by a grant from the United States Department of Homeland

## Abstract

*Burkholderia pseudomallei* is the environmental bacillus that causes melioidosis; a disease clinically significant in Australia and Southeast Asia but emerging in tropical and sub-tropical regions around the globe. Previous studies have placed the ancestral population of the organism in Australia with a single lineage disseminated to Southeast Asia. We have previously characterized *B. pseudomallei* isolates from New Guinea and the Torres Strait archipelago; remote regions that share paleogeographic ties with Australia. These studies identified regional biogeographical boundaries. In this study, we utilize whole-genome sequencing to reconstruct ancient evolutionary relationships and ascertain correlations between paleogeography and present-day distribution of this bacterium in Australasia. Our results indicate that *B. pseudomallei* from New Guinea fall into a single clade within the Australian population. Furthermore, clades from New Guinea are region-specific; an observation possibly linked to limited recent anthropogenic influence in comparison to mainland Australia and Southeast Asia. Isolates from the Torres Strait archipelago were distinct yet scattered among those from mainland Australia. These results provide evidence that the New Guinean and Torres Strait lineages may be remnants of an ancient portion of the Australian population. Rising sea levels isolated New Guinea and the Torres Strait Islands from each other and the Australian mainland, and may have allowed long-term isolated evolution of these lineages, providing support for a theory of microbial biogeography congruent with that of macro flora and fauna. Moreover, these findings indicate that contemporary microbial biogeography theories should consider recent and ongoing impacts of globalisation and human activity.

Security (HSHQDC-10-C-00139). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

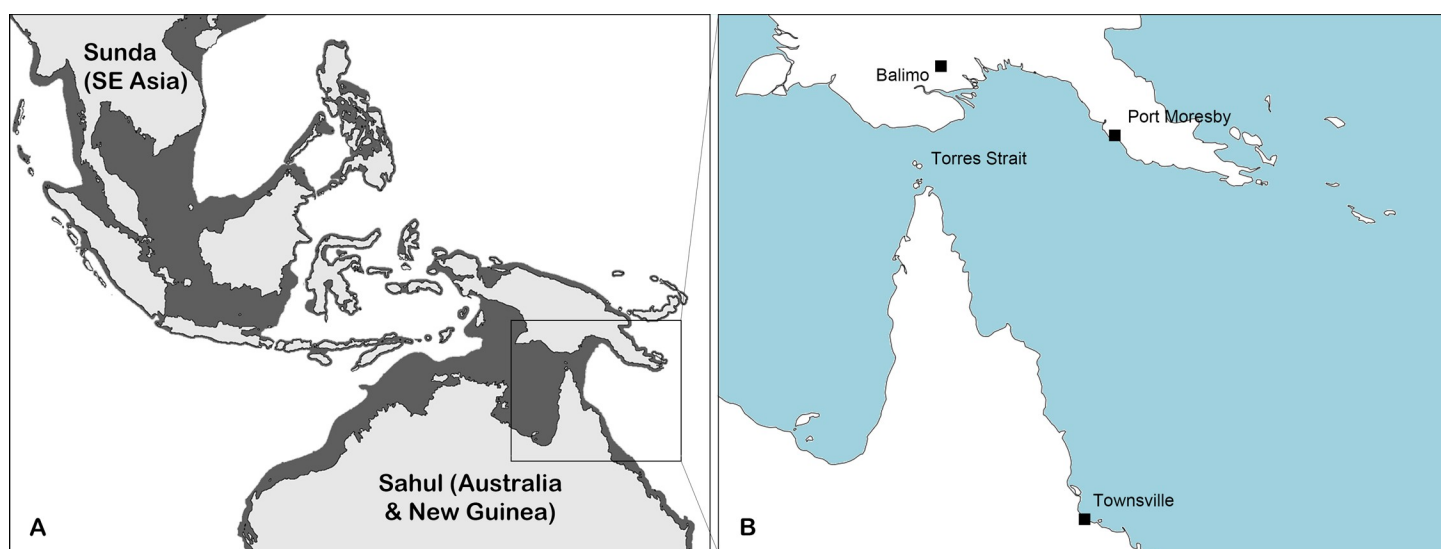
**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

Melioidosis is a severe infection caused by the environmental saprophyte *Burkholderia pseudomallei* and is considered an emerging disease threat throughout tropical and sub-tropical regions globally. In endemic regions the bacterium can be isolated from soil and various freshwater sources [1–4]. Isolates of *B. pseudomallei* from these regions have been extensively characterized using multi-locus sequence typing (MLST), which has demonstrated that mainland Australian isolates are distinct to those from Asia [5–8]. Furthermore, whole genome based phylogenies of Australian isolates in comparison to those from Southeast Asia indicate that the organism most likely originated in Australia before being introduced into Southeast Asia in a single or limited introductory event [7, 9–11]. High genetic diversity in both Australian and Southeast Asian isolates indicate that this dispersion event took place long before present, yet there is very little evidence of admixture between populations. These observations indicate that biogeographical boundaries inhibiting both regional and long-range dispersal of the organism are strong.

The Wallace Line [12] is a transitional area between Asian and Australian ecozones and has been hypothesized to represent the boundary between Asian and Australian *B. pseudomallei* populations [7, 13]. The boundary is based on a deep-water trench that persisted even during the lower sea levels of past glacial periods when much of mainland Asia was united with Indonesia and the Philippines to form the Sunda continent. Similarly, mainland Australia, the Torres Strait and New Guinea united to form the continent of Sahul which remained separated from Sunda by the deep Java and Philippine oceanic trenches [14, 15] (Fig 1). The degree of population isolation and the extent to which the Wallace Line defines the boundary between Australian and Asian *B. pseudomallei* populations is not well defined as few isolates from Indonesia, the Torres Straits, and New Guinea have been characterized. Addressing this question will improve our understanding of *B. pseudomallei* dispersal mechanisms, the potential role of humans and/or animals in its dissemination, and of geographical boundaries that have prevented gene flow between these populations.

Melioidosis has never been reported from the western half of New Guinea (Papua, Indonesia) and is only seldom reported from the east (Papua New Guinea), most likely as a result of poor



**Fig 1. Map of Australasia and Southeast Asia.** a) 21,500 years ago during the last glacial maximum. The shaded regions represent what was dry land during the period. Note that Australia and PNG comprised a single continent (Sahul) and that most of Southeast Asia (Sunda) was linked by land bridges. b) Present day locations of significance for this study. The Torres Strait archipelago is comprised of some 274 individual islands scattered between New Guinea and mainland Australia.

<https://doi.org/10.1371/journal.pone.0206845.g001>

primary healthcare in regional areas [16]. Although it is likely that the organism is widespread throughout lowland regions of the island, only two regions of melioidosis endemicity have been documented: Balimo in the Western Province and Port Moresby in the Central Province. As such, *B. pseudomallei* isolates from the island have only rarely been characterized [17]. We have previously analysed *B. pseudomallei* isolates from the Balimo region of New Guinea, where we identified a low level of genetic diversity [17]. Multi-locus variable-number tandem repeat analysis (MLVA) was used to hypothesize a long-term isolated environmental persistence of *B. pseudomallei* in the region that is subjected to limited anthropological influences [13]. Broader questions regarding the establishment of an isolated *B. pseudomallei* population in the region and their deeper evolutionary origins in a global context remain to be answered. Given the non-random distribution of *B. pseudomallei* genotypes throughout the adjacent Torres Strait archipelago that is indicative of island biogeographical influences and limited dispersal across a marine environment [18], it is likely that the paleogeography of the region is reflected in the present day *B. pseudomallei* population distribution.

Therefore, the aim of this study was to determine the evolutionary relationships among *B. pseudomallei* isolates from New Guinea, the Torres Strait region of northern Queensland and mainland Australia. Moreover, by comparing these isolates to those previously characterised from around the globe, we aim to progress our understanding of the dispersal and biogeography of melioidosis on a global scale.

## Materials and methods

### Ethics statement

Approval and ethical clearance for this study was granted by the Medical Research Advisory Committee (MRAC) of Papua New Guinea under MRAC No 10.03. All clinical isolates collected originate from diagnostic specimens, and as such patients did not provide written informed consent. MRAC is the appropriate body in Papua New Guinea to grant approval for the later use of clinical samples in research.

### Bacterial isolates

This study analysed 12 clinical (from individual patients) and two environmental isolates of *B. pseudomallei* (Table 1), four of which were retrieved from the Balimo region of Papua New Guinea as previously described [17]; five from patients at the Port Moresby Hospital [16, 19], whilst five clinical isolates were from the Torres Strait and the neighbouring region of far northern Queensland [18] (Fig 1). An additional seven isolates from the Townsville region of northern Queensland (TSV17, TSV24, TSV28, TSV30, TSV36, TSV44, TSV51) were sequenced for this study. Whole genome sequences for comparison were sourced online including an additional 19 isolates from mainland Australia, 31 isolates from Southeast Asia and ten isolates from broader regions of the globe. Isolates subject to whole genome sequencing (WGS) from Balimo were chosen based on MLST and MLVA data to include representatives providing the broadest diversity [13]. Remaining isolates were analysed with MLST to ensure that sequenced isolates were selected to best capture existing genetic diversity. Bacteria were stored in the James Cook University and Menzies School of Health Research culture collections at -80°C. Isolates were plated onto Ashdown's agar [20] formulated in house and cultivated at 37°C for 48 hours prior to DNA extraction.

### DNA extraction and qualification

A single, large colony from each plate was removed to 1.5 ml O-ring screw-top microcentrifuge tubes (Sarstedt, Germany) containing 500 µl of Lysis buffer (Corbett Robotics, Australia)

**Table 1. *Burkholderia pseudomallei* isolates from Papua New Guinea and the Torres Strait sequenced for this study.**

ID	Location	Year	Source	MLST	SRA Number
C12	Balimo, Western Province, PNG	2005	Clinical	668	SRR2896252
AG57	Adiba, Western Province, PNG	2001	Environmental	667	SRR2896253
C2	Kimama, Western Province, PNG	1995	Clinical	267	SRR2921951
K41	Kimama, Western Province, PNG	1998	Environmental	267	SRR2896254
MSHR0139	Port Moresby, Central Province, PNG	1987	Clinical	246	SRR2896255
MSHR0141	Port Moresby, Central Province, PNG	1992	Clinical	274	SRR2896256
MSHR1950	Port Moresby, Central Province, PNG	2005	Clinical	340	SRR2896271
POM1	Port Moresby, Central Province, PNG	2002	Clinical	248	SRR2896258
POM2	Port Moresby, Central Province, PNG	2001	Clinical	611	SRR2896259
TSI15	Boigu Island, Australia	2006	Clinical	598	SRR2896260
TSI19	Mabuiag Island, Australia	2002	Clinical	237	SRR2896261
TSI29	Laura, Australia	2004	Clinical	604	SRR2896262
TSI31	Thursday Island, Australia	2005	Clinical	606	SRR2896263
TSI32	Yam Island, Australia	2000	Clinical	610	SRR2896264

<https://doi.org/10.1371/journal.pone.0206845.t001>

and 50 µg of proteinase K (Sigma, Australia). Tubes were incubated at 55°C for two hours in a Hybaid Shake 'n' Stack hybridization oven (Thermo Electron Corporation, MA, USA) with the rotisserie set to the lowest setting. Following lysis, tubes were supplemented with 250 µl of 10M guanidine hydrochloride then passed through Promega Wizard SV Genomic DNA Purification System (Promega, Australia) spin columns with subsequent purification as per the manufacturer's directions. Quality and quantity of DNA was determined by NanoPhotometer (Implen, Germany).

### Multi-locus sequence typing

Isolates were screened for diversity using MLST. Pre-sequencing PCRs were performed in 200 µl thin walled PCR tubes (Sarstedt, Germany) using standard reagents from Promega (Australia) and contained: 1 × GoTaq colourless master mix (Promega), 0.8 µM mixed primers and molecular biology grade H<sub>2</sub>O (Sigma, Australia) to 30 µl. Primers for MLST were as described [21] with the recommended amendments listed on the *B. pseudomallei* MLST website ([pubmlst.net/bpseudomallei](http://pubmlst.net/bpseudomallei)). Cycling conditions consisted of an initial denaturation period of 3 min at 95°C, followed by 40 cycles of 95°C for 30 sec, 62°C for 30 sec and 72°C for 30 sec and a final elongation of 72°C for 10 min. Sequencing products were analysed by electrophoresis using a 1.5% agarose gel to ascertain correct fragment size, concentration and purity against a 100 bp DNA marker (Real Biotech Corporation, Taiwan). Reactions were purified and sequenced by Macrogen (Seoul, South Korea), using ABI PRISM3700 automated sequencing instrumentation (Applied Biosystems, MA, USA). New alleles and sequence types (STs) were submitted to the *Burkholderia pseudomallei* MLST database curator.

### Whole genome sequencing

*Burkholderia pseudomallei* DNA libraries were prepared for multiplexed, paired-end sequencing on the Illumina GAIIx and Illumina HiSeq sequencing platforms (Illumina, USA). Paired-end Illumina whole genome sequence data for each isolate was aligned against both chromosomes of *B. pseudomallei* strain K96243 [22] using BWA-MEM v0.7.5 [23]. Duplicate regions were identified and removed based on a self-alignment of the reference genome using NUCmer v3.23 [24]. Single nucleotide polymorphisms (SNPs) were called by the UnifiedGenotyper

method in GATK v2.7.5 [25, 26] on the binary alignment map (BAM) file [27]. Those SNPs below a minimum depth (10x) or a minimum allele proportion (90%) were removed from subsequent analyses. These methods are wrapped by the Northern Arizona SNP Pipeline (NASP) (<http://tgennorth.github.io/NASP/>) and the In Silico Genotyper pipeline [28]. Sequence data generated during this study have been submitted to the Sequence Read Archive database with accession numbers listed in Table 1.

## Phylogenetic reconstruction

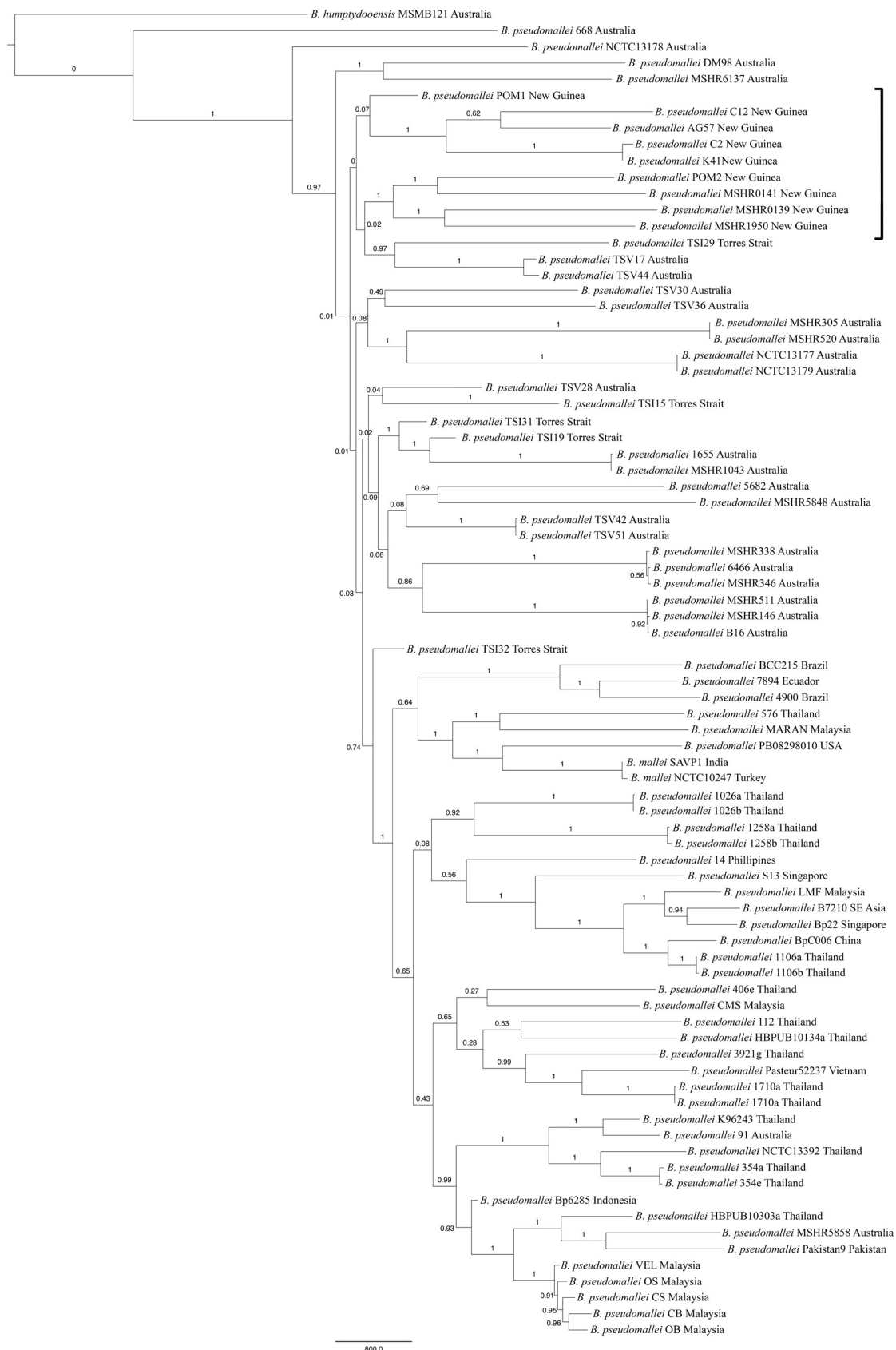
Three algorithms for phylogenetic reconstruction using 49,995 orthologous SNPs were utilized and compared; neighbour joining, maximum parsimony and maximum likelihood. All trees were rooted using the near relative, *Burkholderia humptydooensis* MSMB121. Analyses were performed using MEGA 6.06 [29] with 500 bootstrap replicates for each method. For the neighbour joining phylogeny, evolutionary distances were computed using the maximum composite likelihood method [30]. The model selection tool implemented in MEGA 6.0 [29] was used to determine the best fitting models based on the Bayesian information criterion score. Accordingly, the general time reversible model [31] was used to estimate substitution rates for the maximum likelihood phylogeny, while a discrete gamma distribution was used to model evolutionary variation among sites. Trees were visualized using FigTree version 1.4.2 (Institute of Evolutionary Biology, University of Edinburgh). Differences in tree topologies were calculated with Compare2Trees [32]. A linear regression was performed using TempEst, however a correlation coefficient of -0.26 revealed that a molecular clock signal is not discernible on these isolates.

## Results

MLST resolved 13 sequence types across the New Guinean and Torres Strait isolates examined, with ST 267 shared by isolates C2 and K41, both from the Balimo region of New Guinea. Whole genome sequencing enabled the comparison of an effective core genome size of 1.6MB and represents high quality, non-repetitive bases that led to the discovery of 49,995 SNPs for phylogenetic inference. All three phylogenetic models (neighbour joining, maximum parsimony and maximum likelihood) produced trees that are 79.7–90.0% similar. Despite widespread topological differences, all trees agree, with strong bootstrap support (100%) deeper nodes representing the two distinct populations of *B. pseudomallei* that are geographically separated. Whilst isolates from Australia, New Guinea and the Torres Strait are dispersed throughout the phylograms in a paraphyletic manner, all isolates from Southeast Asia and other regions of the globe, and Australian isolates (BP91 and MSHR5858) form a single monophyletic lineage within the diverse Australian population (Fig 2). Within these major clades there are substantial homoplasies (likely a result of recombination) as indicated by low bootstrap values and a parsimony informative consistency index of 0.32 for the maximum parsimony tree.

All three models indicate that isolates from New Guinea comprise two geographically distinct clades within the Australian population with common ancestry. All isolates from Balimo (e.g., C12, AG57, C2, K41) form a single clade with excellent bootstrap support across all models. Four isolates from Port Moresby (e.g., POM2, MSHR0141, MSHR0139, MSHR1950) form the second New Guinea clade, that was supported by 100% of bootstrap iterations across all methods. The fifth Port Moresby isolate (POM1) is sister to the Balimo clade in the MP tree; however, bootstrap support for this position is poor. In contrast, the neighbour joining and maximum likelihood models suggest that this isolate is the sole representative of yet another New Guinean lineage. All trees indicate that the New Guinean clades stem from basal positions in the phylogeny; however, their exact positions differ based on phylogenetic methodology.





**Fig 2. Maximum parsimony tree of *Burkholderia pseudomallei* isolates constructed using single nucleotide polymorphisms extracted from whole genome sequence data.** This reconstruction demonstrates the relatedness of New Guinean and Torres Strait isolates to other *B. pseudomallei* isolates from around the globe. Bootstrap values (based on 500 replicates) are shown on relevant branches. New Guinean isolates are bracketed to the right of the diagram.

<https://doi.org/10.1371/journal.pone.0206845.g002>

Likewise, the five genomes from the Torres Strait are also placed throughout the more basal parts of the tree and are not closely related to any of the other strains or in well-supported, geographic-specific clades. The exception to this is an isolate from Boigu Island (TSI15) which falls into a clade with isolates from Balimo with 100% bootstrap support in the neighbour joining and maximum likelihood trees. The maximum parsimony tree (Fig 2) shows an alternate placement with an isolate from Townsville, but with very low statistical support (4%). Again, the exact phylogenetic position and topology of many of these deep clades differ between trees and are generally not supported by high bootstrap values. Whilst bootstrap support is 100% for the major clades of Australasia and Asia, topological uncertainty within the Australasian clade is characterized by a wide range of bootstrap support for many intermediate groupings. This correlates with the findings of Pearson et al. (2009) who found that the basal topology of the *B. pseudomallei* tree is marked with uncertainty, probably due to rapid radiation events and ancient lateral gene transfer.

## Discussion

Prior studies [7, 10, 11] have used whole genome sequencing to confirm the presence of two distinct populations of *B. pseudomallei* broadly corresponding to mainland Australia and Southeast Asia, with the Australian population identified as the most ancestral population in all cases. The scarcity of genotypic or phylogenetic overlap from one population (Australasia) to the other (Southeast Asia) is a testament to the strength of ancient oceanic biogeographical barriers in limiting *B. pseudomallei* dissemination as previously hypothesised [7, 18].

We therefore hypothesized that pelagic regions separating other islands would similarly serve as more recent dispersal barriers for *B. pseudomallei* populations in the Torres Strait and New Guinea. The genetic diversity of these isolates as determined using MLST [18] and multiple monophyletic clades (from WGS analyses) from New Guinea and Torres Strait Islands suggest that multiple independent dispersal events would be needed to explain the current distribution. Rather, this distribution is most parsimoniously explained as remnants of an ancient Sahul population. Although the ancient route of *B. pseudomallei* dissemination across Australasia remains incompletely understood, our findings suggest that these populations may have evolved in allopatry when rising sea levels at the end of the Last Glacial Maximum approximately 19,000 years before present [14] isolated them from mainland Australia. Divergence estimates may bolster such hypothesis, however attempts to date using a linear regression model is not likely to yield informative data. This hypothesis does not exclude the possibility that limited dispersal events across pelagic boundaries also occurred, although there is currently insufficient data to support this argument.

Whilst some deep evolutionary nodes in the *B. pseudomallei* phylogeny are well supported among trees, others are particularly difficult to resolve, probably due to both rapid radiations as well as lateral gene transfer events. The limited population structuring among *B. pseudomallei* isolates on a local scale in New Guinea is likely a result of high levels of genetic admixture [7, 8, 33], but not necessarily widespread dispersal. For example, a population gradient, coupled with infrequent dispersal may also produce the population genetic structure seen in this species. Barriers to dispersal are not limited to pelagic boundaries and may also exist, albeit to a lesser extent, across the terrestrial environment. For example, a clonal *B. pseudomallei*

population thought to have been introduced circa 1966 into southwestern Australia has persisted without additional introductions [34, 35]. Other terrestrially linked populations may have remained in isolation for much longer periods. Abnormally high rainfall in central Australia, a region previously not considered endemic for melioidosis, led to the discovery of novel and diverse lineages of *B. pseudomallei* not seen elsewhere in Australia or globally [36]. Even within the highly endemic region of northern Australia, McRobb et al. (2014) demonstrated the existence of geographic structuring between isolates from the Northern Territory and Queensland. Recent analysis of the distribution of STs in the Northern Territory of Australia has also supported limited dispersal of the large number of STs identified to date in that large region, with *B. pseudomallei* spread generally limited to a maximum distance of 45 km [37], although there is one notable exception [38]. In New Guinea, isolates from Balimo and Port Moresby fall into separate monophyletic clades, suggesting that like these other examples, dispersal and subsequent ecological establishment may be uncommon. The frequency of successful introduction events leading to ecological establishment is likely to be variable and highly dependent on geographic boundaries, modes and frequency of dispersal, proximity to endemic regions, anthropogenic factors and ecological conditions.

Human-mediated dispersal has played a critical role in the distribution of many bacterial pathogens [39–42] and continues to do so [43, 44]. There is ample evidence of anthropogenic introductions of *B. pseudomallei* into non-endemic regions of the world [10, 34, 45–47]. It is therefore likely that due to globalisation, human activities are now more important in influencing the population genetics of this species [48]. Human mobility and environmental impact in New Guinea is limited in comparison to other regions, and is reflected in a high diversity of culture, language and human genetics [49]. Given that the only modern modes of human transport between Balimo and Port Moresby are by air or sea, the dissemination of *B. pseudomallei* isolates between the regions by human activities is particularly limited. It is therefore likely that ancient *B. pseudomallei* populations have persisted in New Guinea relatively free of the influence of recent human trade and cultural practices. The observation that isolates from New Guinea form distinct regional clades, whereas those from mainland Australia and Southeast Asia do not conform to such constraints [2, 5], indicates that anthropogenic influences or other mechanisms facilitating dispersal of the organism across a terrestrial environment occur more frequently in other parts of Australia and Asia. Such observation may be linked to the stark contrast between nomadic lifestyles of the traditional mainland Australian people versus the relatively sedentary lifestyles of the diverse human populations in New Guinea.

The limited diversity of *B. pseudomallei* from the island of New Guinea in comparison to that of other regions raises questions as to role of recent anthropogenic influences on the distribution and establishment of new *B. pseudomallei* populations in this region, and by extension the globe. Furthermore, the observations of this study are congruent with contemporary hypotheses that longer-term separation of *B. pseudomallei* populations each side of the Wallace line are responsible for divergent evolution of Southeast Asian and Australasian populations of *B. pseudomallei* in a manner congruent with that of macro flora and fauna in the region. Robust theories of microbial biogeography therefore, must consider not only traditional biogeographical factors on geological time-scales, but account for the recent and ongoing impact of globalisation and human activity.

## Acknowledgments

The authors would like to acknowledge the assistance of John McBride (Cairns Hospital), and Robert Norton (Townsville Hospital), Karina Kennedy (Canberra Hospital) and Diana Martin (ESR Communicable Disease Centre, Porirua, New Zealand) for the provision of bacterial



isolates. This work was supported by a grant from the United States Department of Homeland Security (HSHQDC-10-C-00139).

## Author Contributions

**Conceptualization:** Anthony L. Baker, Talima Pearson, Bart J. Currie, Jeffrey Warner.

**Data curation:** Anthony L. Baker, Talima Pearson, Jason W. Sahl, Crystal Hepp, Erin P. Price, Derek S. Sarovich, Apichai Tuanyok.

**Formal analysis:** Anthony L. Baker, Talima Pearson, Jason W. Sahl, Crystal Hepp, Erin P. Price, Derek S. Sarovich, Apichai Tuanyok.

**Funding acquisition:** Talima Pearson, Paul Keim, Jeffrey Warner.

**Investigation:** Anthony L. Baker, Crystal Hepp, Erin P. Price, Mark Mayo, Jeffrey Warner.

**Methodology:** Anthony L. Baker, Talima Pearson, Jason W. Sahl, Erin P. Price, Derek S. Sarovich, Mark Mayo, Bart J. Currie, Jeffrey Warner.

**Project administration:** Anthony L. Baker, Paul Keim, Jeffrey Warner.

**Resources:** Anthony L. Baker, Talima Pearson, Mark Mayo, Bart J. Currie, Paul Keim, Jeffrey Warner.

**Software:** Anthony L. Baker, Talima Pearson, Jason W. Sahl, Crystal Hepp, Erin P. Price, Derek S. Sarovich.

**Supervision:** Jeffrey Warner.

**Validation:** Anthony L. Baker.

**Visualization:** Anthony L. Baker.

**Writing – original draft:** Anthony L. Baker, Talima Pearson, Jeffrey Warner.

**Writing – review & editing:** Anthony L. Baker, Talima Pearson, Jason W. Sahl, Crystal Hepp, Erin P. Price, Derek S. Sarovich, Mark Mayo, Apichai Tuanyok, Bart J. Currie, Paul Keim, Jeffrey Warner.

## References

1. Yap EH, Thong TW, Tan AL, Yeo M, Tan HC, Loh H, et al. Comparison of *Pseudomonas pseudomallei* from humans, animals, soil and water by restriction endonuclease analysis. *Singapore Med J.* 1995; 36(1):60–2. PMID: [7570138](#)
2. Baker A, Tahani D, Gardiner C, Bristow KL, Greenhill AR, Warner J. Groundwater Seeps Facilitate Exposure to *Burkholderia pseudomallei*. *Appl Environ Microbiol.* 2011; 77(20):7243–6. <https://doi.org/10.1128/AEM.05048-11> PMID: [21873480](#)
3. Mayo M, Kaesti M, Harrington G, Cheng AC, Ward L, Karp D, et al. *Burkholderia pseudomallei* in Unchlorinated Domestic Bore Water, Tropical Northern Australia. *Emerg Infect Dis.* 2011; 17(7):1283–5. <https://doi.org/10.3201/eid1707.100614> PMID: [21762588](#)
4. Baker AL, Warner JM. *Burkholderia pseudomallei* is frequently detected in groundwater that discharges to major watercourses in northern Australia. *Folia Microbiologica.* 2016; 61(4):301–5. <https://doi.org/10.1007/s12223-015-0438-3> PMID: [26620184](#)
5. Cheng AC, Godoy D, Mayo M, Gal D, Spratt BG, Currie BJ. Isolates of *Burkholderia pseudomallei* from Northern Australia are distinct by multilocus sequence typing, but strain types do not correlate with clinical presentation. *J Clin Microbiol.* 2004; 42(12):5477–83. <https://doi.org/10.1128/JCM.42.12.5477-5483.2004> PMID: [15583269](#)
6. Vesaratchavest M, Tumapa S, Day NP, Wuthiekanun V, Chierakul W, Holden MT, et al. Nonrandom distribution of *Burkholderia pseudomallei* clones in relation to geographical location and virulence. *J Clin Microbiol.* 2006; 44(7):2553–7. <https://doi.org/10.1128/JCM.00629-06> PMID: [16825379](#)

7. Pearson T, Giffard P, Beckstrom-Sternberg S, Auerbach R, Hornstra H, Tuanyok A, et al. Phylogeographic reconstruction of a bacterial species with high levels of lateral gene transfer. *BMC Biol.* 2009; 7(1):78.
8. Dale J, Price EP, Hornstra H, Busch JD, Mayo M, Godoy D, et al. Epidemiological Tracking and Population Assignment of the Non-Clonal Bacterium, *Burkholderia pseudomallei*. *PLoS Negl Trop Dis.* 2011; 5(12):e1381. <https://doi.org/10.1371/journal.pntd.0001381> PMID: 22180792
9. Sarovich DS, Garin B, De Smet B, Kaestli M, Mayo M, Vandamme P, et al. Phylogenomic Analysis Reveals an Asian Origin for African *Burkholderia pseudomallei* and Further Supports Melioidosis Endemicity in Africa. *mSphere.* 2016; 1.
10. Chewapreecha C, Holden MTG, Vehkala M, Välimäki N, Yang Z, Harris SR, et al. Global and regional dissemination and evolution of *Burkholderia pseudomallei*. *Nature Microbiology.* 2017; 2:16263. <https://doi.org/10.1038/nmicrobiol.2016.263> PMID: 28112723
11. Price EP, Currie BJ, Sarovich DS. Genomic Insights Into the Melioidosis Pathogen, *Burkholderia pseudomallei*. *Current Tropical Medicine Reports.* 2017; 4(3):95–102.
12. Wallace AR. On the Physical Geography of the Malay Archipelago. *Journal of the Royal Geographical Society of London.* 1863; 33:217–34.
13. Baker A, Pearson T, Price EP, Dale J, Keim P, Hornstra H, et al. Molecular Phylogeny of *Burkholderia pseudomallei* from a Remote Region of Papua New Guinea. *PLoS One.* 2011; 6(3):e18343. Epub 2011/04/13. <https://doi.org/10.1371/journal.pone.0018343> PMID: 21483841
14. Yokoyama Y, Lambeck K, De Deckker P, Johnston P, Fifield LK. Timing of the Last Glacial Maximum from observed sea-level minima. *Nature.* 2000; 406(6797):713–6. <https://doi.org/10.1038/35021035> PMID: 10963593
15. Barrows TT, Stone JO, Fifield LK, Cresswell RG. Late Pleistocene Glaciation of the Kosciuszko Massif, Snowy Mountains, Australia. *Quaternary Research.* 2001; 55(2):179–89.
16. Warner JM, Pelowa D, Currie BJ. Melioidosis – An uncommon but also under-recognized cause of pneumonia in Papua New Guinea. *P N G Med J.* 2010; 54(3–4).
17. Warner JM, Pelowa DB, Gal D, Rai G, Mayo M, Currie BJ, et al. The epidemiology of melioidosis in the Balimo region of Papua New Guinea. *Epidemiol Infect.* 2007:1–7.
18. Baker A, Mayo M, Owens L, Burgess G, Norton R, McBride WJ, et al. Biogeography of *Burkholderia pseudomallei* in the Torres Strait Islands of Northern Australia. *J Clin Microbiol.* 2013; 51(8):2520–5. <https://doi.org/10.1128/JCM.00418-13> PMID: 23698533
19. Currie B. Melioidosis in Papua New Guinea: is it less common than in tropical Australia? *Trans R Soc Trop Med Hyg.* 1993; 87(4):417. PMID: 8249067
20. Ashdown LR. An improved screening technique for isolation of *Pseudomonas pseudomallei* from clinical specimens. *Pathology.* 1979; 11(2):293–7. PMID: 460953
21. Godoy D, Randle G, Simpson AJ, Aanensen DM, Pitt TL, Kinoshita R, et al. Multilocus sequence typing and evolutionary relationships among the causative agents of melioidosis and glanders, *Burkholderia pseudomallei* and *Burkholderia mallei*. *J Clin Microbiol.* 2003; 41(5):2068–79. <https://doi.org/10.1128/JCM.41.5.2068-2079.2003> PMID: 12734250
22. Holden MT, Titball RW, Peacock SJ, Cerdano-Tarraga AM, Atkins T, Crossman LC, et al. Genomic plasticity of the causative agent of melioidosis, *Burkholderia pseudomallei*. *Proc Natl Acad Sci U S A.* 2004; 101(39):14240–5. <https://doi.org/10.1073/pnas.0403302101> PMID: 15377794
23. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:1303.3997–3999*, 2013.
24. Delcher AL, Salzberg SL, Phillippy AM. Using MUMmer to identify similar regions in large sequence sets. *Curr Protoc Bioinformatics.* 2003;Chapter 10:Unit 10.3.
25. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010; 20(9):1297–303. <https://doi.org/10.1101/gr.107524.110> PMID: 20644199
26. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011; 43(5):491–8. <https://doi.org/10.1038/ng.806> PMID: 21478889
27. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009; 25(16):2078–9. <https://doi.org/10.1093/bioinformatics/btp352> PMID: 19505943
28. Sahl JW, Beckstrom-Sternberg SM, Babic-Sternberg J, Gillece JD, Hepp CM, Auerbach RK, et al. The In Silico Genotyper (ISG): an open-source pipeline to rapidly identify and annotate nucleotide variants for comparative genomics applications 2015 2015-01-01 00:00:00.

29. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol.* 2013; 30(12):2725–9. <https://doi.org/10.1093/molbev/mst197> PMID: 24132122
30. Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc Natl Acad Sci U S A.* 2004; 101(30):11030–5. <https://doi.org/10.1073/pnas.0404206101> PMID: 15258291
31. Nei M, Kumar S. *Molecular Evolution and Phylogenetics.* New York: Oxford University Press; 2000.
32. Nye TMW, Liò P, Gilks WR. A novel algorithm and web-based tool for comparing two alternative phylogenetic trees. *Bioinformatics.* 2006; 22(1):117–9. <https://doi.org/10.1093/bioinformatics/bti720> PMID: 16234319
33. McRobb E, Kaestli M, Price EP, Sarovich DS, Mayo M, Warner J, et al. Distribution of *Burkholderia pseudomallei* in northern Australia, a land of diversity. *Appl Environ Microbiol.* 2014; 80(11):3463–8. <https://doi.org/10.1128/AEM.00128-14> PMID: 24657869
34. Currie B, Smith-Vaughan H, Golledge C, Buller N, Sriprakash KS, Kemp DJ. *Pseudomonas pseudomallei* isolates collected over 25 years from a non-tropical endemic focus show clonality on the basis of ribotyping. *Epidemiol Infect.* 1994; 113(2):307–12. PMID: 7523158
35. Chapple SNJ, Sarovich DS, Holden MTG, Peacock SJ, Buller N, Golledge C, et al. Whole-genome sequencing of a quarter-century melioidosis outbreak in temperate Australia uncovers a region of low-prevalence endemicity. *Microbial Genomics.* 2016; 2(7)
36. Yip TW, Hewagama S, Mayo M, Price EP, Sarovich DS, Bastian I, et al. Endemic melioidosis in residents of desert region after atypically intense rainfall in central Australia, 2011. *Emerg Infect Dis.* 2015; 21(6):1038–40. <https://doi.org/10.3201/eid2106.141908> PMID: 25988301
37. Chapple SN, Price EP, Sarovich DS, McRobb E, Mayo M, Kaestli M, et al. *Burkholderia pseudomallei* Genotype Distribution in the Northern Territory, Australia. *Am J Trop Med Hyg.* 2016; 94(1):68–72. <https://doi.org/10.4269/ajtmh.15-0627> PMID: 26526925
38. Aziz A, Sarovich DS, Harris TM, Kaestli M, McRobb E, Mayo M, et al. Suspected cases of intracontinental *Burkholderia pseudomallei* sequence type homoplasy resolved using whole-genome sequencing. *Microb Genom.* 2017 Nov; 3(11).
39. Grimm D. Global spread of leprosy tied to human migration. *Science.* 2005; 308(5724):936–7.
40. Kenefic LJ, Pearson T, Okinaka RT, Schupp JM, Wagner DM, Hoffmaster AR, et al. Pre-Columbian origins for North American anthrax. *PLoS One.* 2009; 4(3):e4813. <https://doi.org/10.1371/journal.pone.0004813> PMID: 19283072
41. Comas I, Coscolla M, Luo T, Borrell S, Holt KE, Kato-Maeda M, et al. Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. *Nat Genet.* 2013; 45(10):1176–82. <https://doi.org/10.1038/ng.2744> PMID: 23995134
42. Wagner DM, Klunk J, Harbeck M, Devault A, Waglechner N, Sahl JW, et al. *Yersinia pestis* and the plague of Justinian 541–543 AD: a genomic analysis. *Lancet Infect Dis.* 2014; 14(4):319–26. [https://doi.org/10.1016/S1473-3099\(13\)70323-2](https://doi.org/10.1016/S1473-3099(13)70323-2) PMID: 24480148
43. Marston CK, Allen CA, Beaudry J, Price EP, Wolken SR, Pearson T, et al. Molecular epidemiology of anthrax cases associated with recreational use of animal hides and yarn in the United States. *PLoS One.* 2011; 6(12).
44. Mokrousov I. *Mycobacterium tuberculosis* phylogeography in the context of human migration and pathogen's pathobiology: Insights from Beijing and Ural families. *Tuberculosis (Edinb).* 2015; 95 Suppl 1:S167–76.
45. Mollaret HH. L'affaire du jardin des plantes ou comment la melioidose fit son apparition en France. *Medecine et Maladies Infectieuses.* 1988; 18(Supplement 4):643–54.
46. Ouadah AHE, Mohd. Zahedi D, Perumal R. Animal Melioidosis Surveillance In Sabah. *The Internet Journal of Veterinary Medicine.* 2007; 2(2).
47. Doker TJ, Sharp TM, Rivera-Garcia B, Perez-Padilla J, Benoit TJ, Ellis EM, et al. Contact investigation of melioidosis cases reveals regional endemicity in Puerto Rico. *Clin Infect Dis.* 2015; 60(2):243–50. <https://doi.org/10.1093/cid/ciu764> PMID: 25270646
48. Currie BJ, Kaestli M. Epidemiology: A global picture of melioidosis. *Nature.* 2016; 529(7586):290–1. <https://doi.org/10.1038/529290a> PMID: 26791716
49. Allen B. Human geography of Papua New Guinea. *Journal of Human Evolution.* 1983; 12(1):3–23.